

DATA MIGRATION METHOD

Publication number: JP2001249853

Publication date: 2001-09-14

Inventor: WATANABE NAOKI; TAKAMOTO YOSHIFUMI

Applicant: HITACHI LTD

Classification:

- International: G06F13/14; G06F3/06; G06F11/14; G06F12/00;
G06F12/16; G11B20/10; G06F13/14; G06F3/06;
G06F11/14; G06F12/00; G06F12/16; G11B20/10;
(IPC1-7): G06F12/16; G06F3/06; G06F12/00;
G06F13/14; G11B20/10

- European: G06F3/06M; G06F11/14A4B1M6

Application number: JP20000063289 20000303

Priority number(s): JP20000063289 20000303

Also published as:

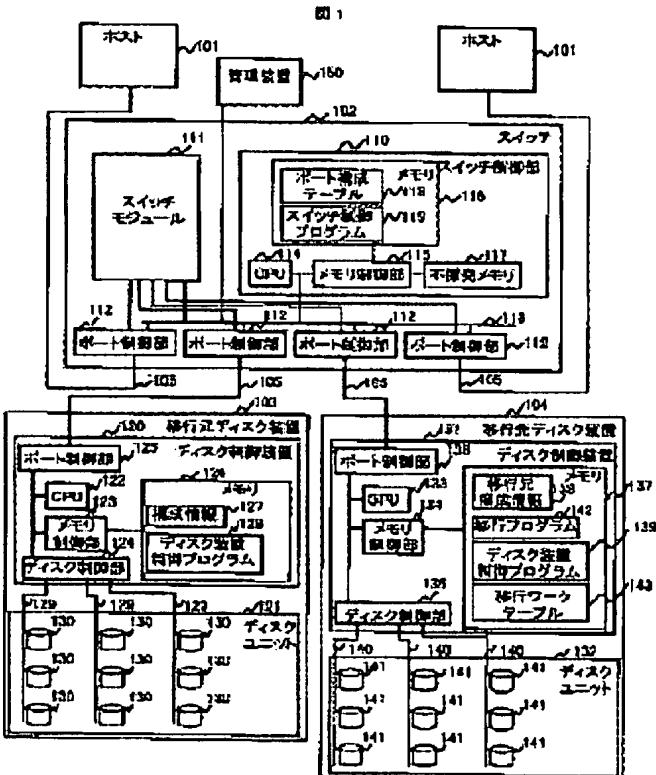
EP1130514 (A2)
US6976103 (B1)

[Report a data error here](#)

Abstract of JP2001249853

PROBLEM TO BE SOLVED: To provide a data migration method suitable for SAN(Storage Area Network) environment and its device.

SOLUTION: A disk device at migration destination is connected with a switch in which a host is connected with a disk device at a migration origin. At this memory, the same value as that of a port ID provided to an F-Port of the switch is provided to a port ID of the disk device at the migration destination and the disk device at the migration destination is made not to be recognized by the host 101. The disk information at the migration destination reads the constitutional information on the disk device at the migration origin. After that, a physical port ID 602 of the disk device 103 at the migration origin is replaced with the physical port ID 602 of the disk device 104 at the migration destination in the switch. The disk device at the migration destination constructs volume according to a logical volume number on the disk device at the migration original and the size of the volume and stores data in the disk device at the migration origin in corresponding volume.



Data supplied from the **esp@cenet** database - Worldwide

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2001-249853

(P2001-249853A)

(43)公開日 平成13年9月14日 (2001.9.14)

(51)Int.Cl'	識別記号	F I	テ-71-ト*(参考)
G 0 6 F 12/16	3 1 0	G 0 6 F 12/16	3 1 0 M 5 B 0 1 4
3/06	3 0 4	9/06	3 0 4 F 5 B 0 1 8
12/00	5 0 1	12/00	5 0 1 B 5 B 0 6 5
	5 1 1		5 1 1 Z 6 B 0 8 2
19/14	3 1 0	13/14	3 1 0 F 5 D 0 4 4

審査請求 未請求 請求項の数12 OL (全24頁) 最終頁に続く

(21)出願番号 特願2000-63289(P2000-63289)

(22)出願日 平成12年3月3日 (2000.3.3)

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 渡邊 直企

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(72)発明者 高本 良史

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(74)代理人 100075096

弁理士 作田 康夫

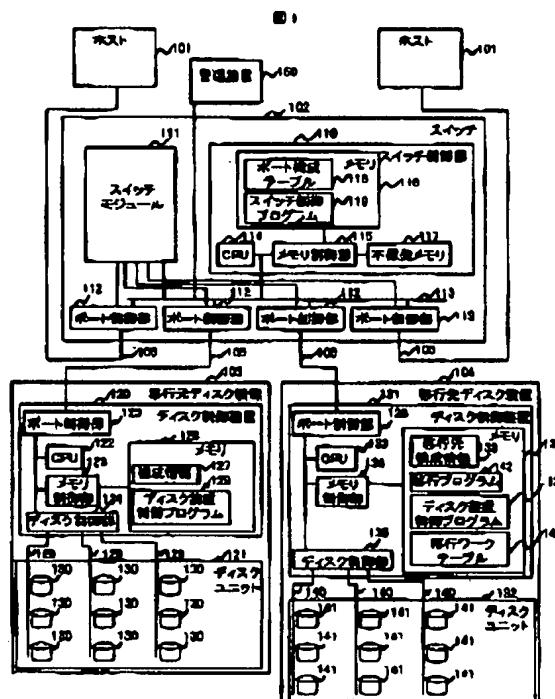
最終頁に続く

(54)【発明の名称】 データ移行方法

(57)【要約】

【課題】 SAN (Storage Area Network) 環境に好適なデータ移行方法、及びその装置を提供することである。

【解決手段】 ホストと移行元ディスク装置が接続されているスイッチに、移行先ディスク装置を接続する。この時点では、移行先ディスク装置のポートIDに、スイッチのF_Portに与えられたポートIDと同じ値を与えておく、移行先ディスク装置はホスト101から認識されないようにしておく。移行先ディスク情報は移行元ディスク装置の構成情報を読み出す。その後、スイッチにおいて、移行元ディスク装置103の物理ポートID 602と移行先ディスク装置104の物理ポートID 602とを入れ替える。移行先ディスク装置は、移行元ディスク装置上の論理ボリューム番号とそのボリュームの大きさに応じたボリュームを構築し、移行元ディスク装置内のデータを対応するボリュームに格納する。



(2)

特開2001-249853

2

【特許請求の範囲】

【請求項1】スイッチと通信チャネルにより接続され、上記スイッチと通信チャネルにより接続されるホストコンピュータからのリード／ライト要求を受け付ける第1のストレージサブシステムから、第2のストレージサブシステムにデータを移行する方法であって、上記第2のストレージサブシステムを通信チャネルにより上記スイッチと接続し、上記スイッチを介して、上記第1のストレージサブシステムに格納されているデータを上記第2のストレージサブシステムに書き込むことを特徴とするデータ移行方法。

【請求項2】スイッチと通信チャネルにより接続され、上記スイッチと通信チャネルにより接続されるホストコンピュータからのリード／ライト要求を受け付ける第1のストレージサブシステムから、第2のストレージサブシステムにデータを移行する方法であって、上記第2のストレージサブシステムを通信チャネルにより上記スイッチと接続し、上記第1のストレージサブシステムに構築されている論理ボリュームの数、及びそれらのサイズを、上記スイッチを介して、上記第2のストレージサブシステムに読み出し、上記第1のストレージサブシステムに構築されている論理ボリュームと同数及び同サイズの論理ボリュームを上記2のストレージサブシステムに構築し、ボリューム単位で、上記第1のストレージサブシステムに格納されているデータを、上記スイッチを介して上記第2のストレージサブシステムに書き込むことを特徴とするデータ移行方法。

【請求項3】それぞれポート識別子が割り当てられている複数のポートを有するスイッチの一つのポートと通信チャネルにより接続される第1のストレージサブシステムであって、上記スイッチの他のポートと通信チャネルにより接続されるホストコンピュータから、上記スイッチを介してリード／ライト要求を受け付ける上記第1のストレージサブシステムから、第2のストレージにデータを移行する方法であって、上記第2のストレージサブシステムを、上記スイッチの上記第1のディスク装置及び上記ホストコンピュータが接続されているポート以外のポートに通信チャネルにより接続し、上記第1のストレージサブシステムが接続されているポートに割り当てられているポート識別子と、上記第2のストレージサブシステムが接続されているポートに割り当てられているポート識別子とを互いに入れ替え、上記第1のストレージサブシステム内のデータを、上記スイッチを介して上記第2のストレージサブシステムに書き込み、上記第1のストレージサブシステムから上記第2のストレージサブシステムへのデータの書き込みが完了したか否かをボリューム単位で管理し、上記ホストコンピュータから上記第1のストレージサブシステムに対するリード又はライト要求があった場合、その要求は上記第2のストレージサブシステムに送信され、そのリード又はライト要求があったデータが既に上記第2のストレージサブシステムに書き込まれている場合には、上記第2のストレージサブシステムにより、その書き込

まれているデータに対してリード又はライト処理が行われ、そのリード又はライト要求があったデータがまだ上記第2のストレージサブシステムに書き込まれていない場合には、そのデータが上記第2のストレージサブシステムに書き込まれ、上記第2のストレージサブシステムにより、その書き込まれているデータに対してリード又はライト処理が行われることを特徴とするデータ移行方法。

【請求項4】請求項3に記載のデータ移行方法であって、上記第1のストレージサブシステムと上記第2のストレージサブシステムとの間を通信チャネルにより直接接続し、上記第1のストレージサブシステム内のデータを、上記スイッチを介してではなく、上記第1のストレージサブシステムと上記第2のストレージサブシステムとの間を直接接続する通信チャネルを介して、上記第2のストレージサブシステムに書き込むことを特徴とするデータ移行方法。

【請求項5】それぞれポート識別子が割り当てられている複数のポートを有するスイッチの一つのポートと通信チャネルにより接続される第1のストレージサブシステムであって、上記スイッチの他のポートと通信チャネルにより接続されるホストコンピュータから、上記スイッチを介してリード／ライト要求を受け付ける上記第1のストレージサブシステムから、第2のストレージにデータを移行する方法であって、上記第2のストレージサブシステムを、上記スイッチの上記第1のディスク装置及び上記ホストコンピュータが接続されているポート以外のポートに通信チャネルにより接続し、上記第1のストレージサブシステムが接続されているポートに割り当てられているポート識別子と、上記第2のストレージサブシステムが接続されているポートに割り当てられているポート識別子とを互いに入れ替え、上記第1のストレージサブシステムに構築されている論理ボリューム情報を読み出し、上記第2のストレージサブシステムに上記第1のストレージサブシステムに構築されている論理ボリュームと同数及び同サイズの論理ボリュームを構築し、上記第1のストレージサブシステム内のデータを、上記スイッチを介して上記第2のストレージサブシステムに書き込み、上記第1のストレージサブシステムから上記第2のストレージサブシステムへのデータの書き込みが完了したか否かをボリューム単位で管理し、上記ホストコンピュータから上記第1のストレージサブシステムに対するリード又はライト要求があった場合、その要求は上記第2のストレージサブシステムに送信され、そのリード又はライト要求があったデータが既に上記第2のストレージサブシステムへの書き込みが完了している場合には、上記第2のストレージサブシステムにより、その書き込まれているデータに対してリード又はライト処理が行われ、そのリード又はライト要求があったデータがまだ上記第2のストレージサブシステムへの書き込みが

(3)

特開2001-249853

3

完了していない場合には、そのデータが上記第2のストレージサブシステムに書き込まれ、上記第2のストレージサブシステムにより、その書き込まれているデータに対してリード又はライト処理が行われることを特徴とするデータ移行方法。

【請求項6】 請求項5に記載のデータ移行方法であつて、上記第2のストレージサブシステムが、上記スイッチを介して上記第1のストレージサブシステムから上記論理ボリューム情報を読み出し、上記第2のストレージサブシステムへのデータの書き込みが完了したか否かをボリューム単位で管理することを特徴とするデータ移行方法。

【請求項7】 請求項5に記載のデータ移行方法であつて、上記スイッチが、上記第1のストレージサブシステムから上記論理ボリューム情報を読み出し、上記第2のストレージサブシステムに上記第1のストレージサブシステムに構築されている論理ボリュームと同数及び同サイズの論理ボリュームを構築し、上記第2のストレージサブシステムへのデータの書き込みが完了したか否かをボリューム単位で管理することを特徴とするデータ移行方法。

【請求項8】 請求項5に記載のデータ移行方法であつて、上記スイッチに接続された情報処理装置が、上記第1のストレージサブシステムから上記論理ボリューム情報を読み出し、上記第2のストレージサブシステムに上記第1のストレージサブシステムに構築されている論理ボリュームと同数及び同サイズの論理ボリュームを構築し、上記第2のストレージサブシステムへのデータの書き込みが完了したか否かをボリューム単位で管理することを特徴とするデータ移行方法。

【請求項9】 それぞれポート識別子が割り当てられている複数のポートを有するスイッチの一つのポートと通信チャネルにより接続される第1のストレージサブシステムであつて、上記スイッチの他のポートと通信チャネルにより接続されるホストコンピュータから、上記スイッチを介してリード/ライト要求を受け付ける上記第1のストレージサブシステムから、第2のストレージにデータを移行する方法であつて、上記第2のストレージサブシステムを、上記スイッチの上記第1のディスク装置及び上記ホストコンピュータが接続されているポート以外のポートに通信チャネルにより接続し、上記ホストコンピュータからのリード/ライト要求先を上記第1のストレージサブシステムから上記第2のストレージサブシステムに切り替え、上記スイッチを介して上記第2のストレージサブシステム内のデータを、上記スイッチに書き込み、上記ホストコンピュータからリード又はライト要求があったデータが既に上記第2のストレージサブシステムに書き込まれている場合には、上記第2のストレージサブシステムにより、その書き込まれているデータに対してリード又はライト処理が

10

20

30

40

50

4

行われ、そのリード又はライト要求があったデータがまだ上記第2のストレージサブシステムに書き込まれていない場合には、そのデータが上記第2のストレージサブシステムに書き込まれ、上記第2のストレージサブシステムにより、その書き込まれているデータに対してリード又はライト処理が行われることを特徴とするデータ移行方法。

【請求項10】 請求項9に記載のデータ移行方法であつて、上記第1のストレージサブシステムに構築されている論理ボリュームの数、及びそれらのサイズを、上記スイッチを介して、上記第2のストレージサブシステムに読み出し、上記第1のストレージサブシステムに構築されている論理ボリュームと同数及び同サイズの論理ボリュームを上記2のストレージサブシステムに構築し、ボリューム単位で、上記第1のストレージサブシステムに格納されているデータを、上記スイッチを介して上記第2のストレージサブシステムに書き込むことを特徴とするデータ移行方法。

【請求項11】 請求項10に記載のデータ移行方法であつて、上記第2のストレージサブシステムは、上記第2のストレージサブシステムへのデータの書き込みが完了したか否かをボリューム単位で管理することを特徴とするデータ移行方法。

【請求項12】 請求項1乃至請求項11の何れかに記載のデータ移行方法であつて、上記通信チャネルはファイバチャネルであることを特徴とするデータ移行方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、ストレージエリアネットワーク（以下「SAN」という。）に新ディスク装置（移行先ディスク装置）を追加接続し、SANに既に接続されている旧ディスク装置（移行元ディスク装置）からその新ディスク装置にデータを移行する方法及びその装置に関する。

【0002】

【従来の技術】 従来のコンピュータシステムでは、複数のサーバがLAN等のネットワークで接続され、各コンピュータには、それぞれ、ディスク装置が直接接続される構成となっていた。各ディスク装置内のデータは、それが直接接続されるサーバにより管理されていた。つまり、データは分散して管理されていた。

【0003】しかし、最近では、複数のサーバ、複数のディスク装置、バックアップ装置等がスイッチ、ハブ等で接続される構成をとるSANがホットなトピックとなっている。これらの装置とスイッチ又はハブ等との間の物理的な接続には、ファイバチャネルが用いられる。SANを用いてシステムを構成するメリットは、スケーラビリティに優れていること、複数のディスク装置に分散したデータを統合して一元管理を行えるため、管理コストの低減を実現できる、といった点にある。そのため、SAN

(4)

特開2001-249853

5

は、大規模システムの構築に適している。SANでは、多くの資源が相互接続されるため、それらの資源の管理が重要な議題となる。例えば、ディスク装置の管理としては、論理ボリュームの管理（バックアップ等）が挙げられる。現在は、それらの管理方法が、SNIA (Storage Networking Industry Association) , NSIC (National Storage Industry Consortium) において検討されている。

【0004】一方、米国特許第5,680,640号には、第1のデータストレージシステムから第2のデータストレージシステムへデータを移行するシステム及びその方法が開示されている。その概要は次の通りである。ホスト、ネットワーク、又は他のデータ処理システムにあらかじめ接続されている第1のデータストレージシステムが、上記ホストから切り離され、第2のデータストレージシステムに接続される。上記第2のデータストレージシステムは、上記ホスト、又は上記データ処理システムに接続される。上記第2のデータストレージシステムは、どのデータ要素が上記第2のストレージシステムに格納されているか、そして、上記第1のデータストレージシステム上のどのデータが上記第2のストレージシステムにコピーされているかを示すデータマップ、又はデータテーブルを有している。上記ホスト、上記ネットワーク、又は上記データ処理システムが上記第2のストレージシステムにデータを要求すると、上記第2のデータストレージシステムは、そのデータが上記第2のデータストレージシステム、又は上記第1のストレージシステムに格納されているか否かを判断する。もし、データが上記第2のデータストレージシステムに格納されれば、上記ホスト等に対して、そのデータを使用可能にする。もし、そのデータが上記第2のデータストレージシステムに格納されていない場合には、上記第2のデータストレージシステムは上記第1のデータストレージシステムに要求を発行し、上記ホスト等に対して、そのデータを使用可能にする。そして、上記第2のデータストレージシステムは、そのデータを自分のシステムに書き込み、上記データマップ又は上記データテーブルを更新する。上記第2のデータストレージシステムはビジーでないとき、上記データマップ又は上記データテーブルをスキヤンし、上記第1のデータストレージシステムのどのデータが、まだ自分のシステムにコピーされていないかを判断し、そのデータのコピーと、上記データマップ又は上記データテーブルの更新を実行する。

【0005】

【発明が解決しようとする課題】米国特許第5,680,640号に開示されるデータ移行システム及びその方法によれば、上記ホスト等と独立して、上記第1のデータストレージシステムから上記第2のデータストレージシステムへのデータ移行を実行することができる。

【0006】しかし、上記データ移行システムおよびそ

10

6

の方法では、上記第2のデータストレージシステムの導入時に、上記第1のデータストレージシステムを上記ホストから切り離し、上記第2のデータストレージシステムに接続し、上記第2のデータストレージシステムを上記ホストに接続するという手順が必要となる。そのため、少なくとも上記第1のデータストレージシステムが上記ホストから切り離されてから、上記第2のデータストレージシステムが上記ホスト及び上記第1のストレージシステムに接続されるまでの間は、上記ホストはI/O (Input/Output) 要求を発行することができない。また、その間、上記第1のデータストレージシステムを使用する上記ホスト上のアプリケーション等も一時的に停止される必要がある。データ移行に伴うコストをより抑えるためには、そのI/Oを発行できない時間、アプリケーションを停止する時間をより短縮する必要がある。

20

【0007】また、上記データ移行システムおよびその方法では、上記第1のデータストレージシステムと上記第2のデータストレージシステムとを接続するために、それらのデータストレージシステムの各々に、チャネルを新たに設けなければならない。

20

【0008】さらに、米国特許第5,680,640号には、SAN環境下のデータ移行システム及びその方法が開示されていない。

【0009】本発明の目的は、SAN環境に適したデータ移行システム及びその方法を提供することにある。

【0010】

【課題を解決するための手段】本発明の代表的なデータ移行の概要は次の通りである。

30

【0011】ホストコンピュータと第1のディスク装置とがそれぞれ通信チャネルによりスイッチのポートに接続されており、上記第1のディスク装置は上記スイッチを介して上記ホストコンピュータからリード/ライト要求を受け付けているものとする。上記スイッチの各ポートには物理ポートIDと論理ポートIDとが割り当てられており、上記スイッチは、物理ポートIDと論理ポートIDとの対応関係を保持するテーブルを有している。第2のディスク装置を上記ホストコンピュータが接続されているポート及び上記第1のディスク装置が接続されているポート以外の上記スイッチのポートに接続する。上記第2のディスク装置は、上記スイッチを介して、上記第1のディスク装置の構成情報（例えば、論理ボリューム数、各論理ボリュームのサイズ）を取得する。その後、上記第1のディスク装置が接続されるスイッチのポートに割り当てられている物理ポートIDと論理ポートIDとの対応関係と、上記第2のディスク装置が接続されるスイッチのポートに割り当てられている物理ポートIDと論理ポートIDとの対応関係とを入れ替える。具体的には、上記第1のディスク装置が接続されるスイッチのポートに割り当てられている論理ポートIDと上記第2のディスク装置が接続されるスイッチのポートに割り当てられて

40

40

50

50

50

50

50

50

50

50

50

50

(5)

7

いる論理ポートIDとを入れ替える。これにより、ホストコンピュータから上記第1のディスク装置にアクセスしても、実際には、上記第2のディスク装置にアクセスされることになる。このポートID切替処理を行った後、上記第2のディスク装置上に、上記第1のディスク上の構成情報に対応する論理ボリュームが構築され、上記第1のディスク装置内のデータが上記第2のディスク装置に移行される。上記ホストコンピュータから、上記第2のディスク装置に既に移行されているデータに対してリード又はライト要求が合った場合には、上記第2のディスク装置によりそのデータに対してその処理が行われる。上記ホストコンピュータから、上記第2のディスク装置にまだ移行されていないデータに対してリード又はライト要求が合った場合には、そのデータが上記第1のディスク装置から上記第2のディスク装置に読み出され、上記第2のディスク装置によりそのデータに対してその処理が行われる。

【0012】その他の本願が提供するデータ移行方法は、発明の実施の形態の欄で明らかにされる。

【0013】

【発明の実施の形態】以下、実施例1乃至実施例5は、SAN環境に好適なデータ移行方法及びその装置を提供する。

【0014】【実施例1】本実施例では、ホストと移行元ディスク装置が接続されているスイッチに、移行先ディスク装置が接続され、そのスイッチを用いて、移行元ディスク装置内のデータが、移行先ディスク装置に移行される。従って、本実施例によれば、ホストと移行元ディスク装置との間の接続を切り離し、移行先ディスク装置をホスト及び移行元ディスクに接続するという作業は必要なく、単に、移行先ディスク装置をスイッチに接続するだけなので、上述の米国特許第5,680,640号に開示されるデータ移行システム及びその方法に比べて、移行先ディスク装置を追加する作業が短縮される。そのため、データ移行に伴うコストもより抑えることができる。また、本実施例によれば、データ移行のための専用チャネルを設ける必要もないので、ハードウェアのコストも、より抑えることができる。

【0015】以下、図1から図10を用いて、本実施例を説明する。

【0016】(1) システム構成

図1は、本発明のデータ移行システム及びその方法が適用されるコンピュータシステムの第1の実施例を説明するための図である。2つのホスト101、移行元ディスク装置103、及び移行先ディスク装置104が、スイッチ102により接続されており、これらがSANを形成している。ホスト101、移行元ディスク装置103、及び移行先ディスク装置104とスイッチ102との間はファイバチャネル105で接続されている。なお、本発明は、移行元ディスク装置103から移行先ディスク装置104へデータ移行するとい

特開2001-249853

8

うものであるから、本発明を実施する上で、ホストの台数は制限されない。

【0017】図5は、ファイバチャネル105において使用されるフレームの構造を示す。フレームはフレームの先頭を示すSOF (Start Of Frame) 501と、終わりを示すEOF (End Of Frame) 505との間にFRAME HEADER 502、DATA FIELD 503とCRC 504 (Cyclic Redundancy Check) が入る。FRAME HEADER 502はフレームの制御情報を含み、受信側アドレスであるD_ID (Destination ID) 507と送信元アドレスであるS_ID (Source ID) 508、ルーティングの制御を行うR_CTL (Routing Control) 506、データの構造を示すTYPE 510、フレームのシーケンスやエクスチェンジの制御を行うF_CTL (Frame Control) 511、送信元と受信側のシーケンスの識別をおこなうSEQ_ID (Sequence_ID) 512、各シーケンス毎のフレームの数のカウント値を示すSEQ_CNT (Sequence count) 514、データフィールドの制御情報であるDF_CTL (Data Field Control) 513を含む。ファイバチャネル105のフレームについてはANSIのX3.230 FC-PH (Fibre Channel Physical and Signaling Interface) に詳しい。

【0018】図3は、図1に示されるコンピュータシステムの論理的なネットワーク接続構成を示す。そのコンピュータシステムでは、ファブリック301は一つのスイッチ102で構成されている。ホスト101、ディスク装置との接続に使用するスイッチ102のポート303は、F (Fabric) ポートと呼ばれる。また、ホスト101、ディスク装置のポート302は、N (Node) ポートと呼ばれる。図4に示すように、ファブリック401は、複数のスイッチ102で構成されることも可能である。スイッチ間を接続するスイッチのポートはE (Expansion) ポート401と呼ばれる。

【0019】図3のようにファブリックが一つのスイッチ102で構成されている場合、ホスト101から移行先ディスク装置104へのデータ転送は次のようにして行われる。ホスト101は、自分のポートのIDをS_ID 508に格納し、フレームの送信先のポートIDをD_ID 507に格納しフレームを送信する。スイッチ102は、フレームを受信すると、D_ID 507に格納されているポートIDを調べ、そのポートIDと一致するF_Portから、移行先ディスク装置104のN_Portへフレームを送信する。S_ID 508及びD_IDは論理的なポートIDであり、また、図4に示したように、ファブリック401が複数のスイッチ102で構成されている場合、ホスト101から移行先ディスク装置104へのデータ転送は次のようにして行われる。ホスト101と直接接続されているスイッチ102は、ホスト101からフレームを受信すると、D_ID 507に格納されているポートIDを調べる。しかし、そのスイッチは、そのポートIDと一致するF_Portを持っていないので、E_Portから他のスイッチにフレームを送信する。そして、そのポートIDと一致するF_Portを持っている移行先ディスク装置104に直接接続されているスイッチが、そのF_Portから移行先ディスク

(6)

9

装置104のN_Portへフレームを送信する。以上、ホスト101から移行先ディスク装置104へのデータ転送を例により説明したが、いずれの装置間のデータ転送も同様に行われる。

【0020】再び、図1を参照して、スイッチ102、移行元ディスク装置103、移行先ディスク装置104、及びホスト101の順で、それらの構成を説明する。

【0021】(A) スイッチ102の構成

スイッチ102は、スイッチ制御部110、スイッチモジュール111、及びポート制御部112とを含む。

【0022】スイッチ制御部110は、スイッチ102内の制御を行うものであり、CPU114、メモリ制御部115、メモリ116、及び不揮発メモリ117とを含む。

【0023】CPU114は、内部バス113を用いて、各制御部と制御情報及びデータのやり取りを行う。

【0024】不揮発メモリ117は、メモリ制御部115によりリード／ライトが制御され、スイッチ102の制御に必要なスイッチ制御プログラム119、及びそのプログラムの実行時に必要なポート構成テーブル118等のデータを格納する。

【0025】図6は、ポート構成テーブル118の構成例を示す図である。図6(A)は、ポート切替処理前のポート構成テーブルを示しており、図6(B)は、そのポート切替処理後のポート構成テーブルを示している。ポート切替処理については後述する。ポート構成テーブル118は、ポートの構成を示すものであり、論理的なポートIDを示す論理ポートID 601、及び物理的なポートIDを示す物理ポートID 602を含む。その他に、ファイバチャネルの転送サービスクラスの規定、ループ等のポートの種別が含まれていてもよい。本実施例では、図20に示すように、論理ポートID_0にホスト101が、論理ポートID_1に移行元ディスク装置103が、論理ポートID_4にホスト101が接続されているものとする。また、論理ポートID_2に移行先ディスク装置104が接続されるものとする。

【0026】メモリ116は、メモリ制御部115によりリード／ライトが制御される。スイッチ制御プログラム119がCPU114により実行されるとき、そのプログラムが不揮発メモリ117から読み出され、メモリ116上に格納される。また、その際、ポート構成テーブル118も、必要に応じてCPU114が不揮発メモリ117から読み出され、メモリ116上に格納される。

【0027】ポート制御部112は、データの符号化／複合化等の制御を行う。具体的には、ポート制御部112は、受信したファイバチャネルのフレームからシリアルデータを取り出してパラレルデータに変換し、また、そのフレームから送信先ID等、スイッチングに必要な情報を取り出してスイッチモジュール111に送信する。また、ポート制御部112は、その逆の処理も行う。

【0028】スイッチモジュール111は複数のポートを有しており、それらは、それぞれ複数のポート制御部11

特開2001-249853

10

2に接続される。スイッチモジュール111は、内部バス113を介してCPU114と接続されており、CPU114により制御される。スイッチモジュール111は、ポート制御部112からデータを受信すると、そのデータの送信先IDに従い、出力すべきポートを切り替え、データを送信する。スイッチモジュール111はクロスバススイッチ等で構成すればよい。

【0029】スイッチ102に接続される制御装置150は、スイッチ102の各種パラメータの設定、本実施例の最も特徴できな機能である移行処理の制御及び情報管理を行う。また、管理装置150は、後述する移行先ディスク装置104内の移行プログラムを起動させるためのコマンド、移行情報（例えば移行中、移行終了、異常発生）を取得するためのコマンドを持っている。

【0030】なお、図1では、管理装置150はスイッチ102に直接接続されているが、図22に示すように、ホスト101、スイッチ102、移行元ディスク装置103、移行先ディスク装置104、及び制御装置150をLAN (Local Area Network) に接続し、Web等により上記の設定を行うようにしてもよい。

【0031】(B) 移行元ディスク装置102及び移行先ディスク装置103の構成

本実施例では、説明を簡単にするために、移行元ディスク装置103と移行先ディスク装置104とは、メモリ上のプログラム及び情報以外は同様の構成とする。

【0032】移行元ディスク装置103は、ディスク制御装置120とディスクユニット121とを含む。

【0033】ディスクユニット121は、複数のディスクドライブ130を含んでいる。これらのディスクドライブ130はインターフェース (I/F) 129 (ファイバチャネル、SCSI等) によりディスク制御部124に接続されている。しかし、本発明を実施する上では、ディスクドライブ130の台数は制限されない。

【0034】ディスク制御装置120は、CPU 122、メモリ制御部123、ポート制御部125、ディスク制御部124、及びメモリ126を含む。CPU 122、メモリ制御部123、ポート制御部125、及びディスク制御部124とは内部バスで接続されている。

【0035】CPU 122は、それらの制御部と、その内部バスを用いて制御情報、データ等のやり取りを行う。CPU 122は、スイッチ102経由でホスト101から送信されたコマンドに応じて、そのコマンドの処理に必要なディスクドライブ130にリード・ライト命令を発行する。CPU 122は、複数のディスクドライブ130により、よく知られているRAID 0～5の構成を構築し、ホスト101に対して論理的なボリュームを提供する。

【0036】ポート制御部125は、ファイバチャネル105により、スイッチ102のポート制御部112に接続されており、その機能はポート制御部112と同様である。

【0037】メモリ126は、メモリ制御部123に接続され

50

(7)

特開2001-249853

11

ており、ディスク装置を制御するディスク装置制御プログラム128、該プログラムの実行時に必要なデータ、及び構成情報127を格納する。構成情報127は後述される。ディスク装置制御プログラム128は、CPU 122により実行され、ポート制御部125及びディスク制御部124の制御、及びホスト101から受信したリード/ライトコマンドの処理を行うプログラムである。

【0038】上述のように、移行先ディスク装置104の構成要素は、メモリ137上のプログラム及び情報以外は、移行元ディスク装置103の構成要素と同様であるので、メモリ137上のプログラム及び情報のみ説明する。メモリ137上には、移行元ディスク装置103のメモリ126に格納されている構成情報127が読み取られ、それが入力されたものである移行元構成情報138と、データの移行時に使用する移行プログラム142、データの移行の状態を示す移行ワークテーブル143が格納される。

【0039】図8は、移行ワークテーブル143の構成例を示す。移行ワークテーブル143は、ボリュームの番号801、スロットの番号802、及び該スロットの状態を表すステータス803を含む。

【0040】(C) ホスト101の構成

図2は、ホスト101の構成を示す。ホスト101は、CPU 201、メモリ制御部202、ポート制御部206、ディスク制御部204、ディスクドライブ205、及びメモリ203などを含む。

【0041】CPU 201、メモリ制御部202、ポート制御部206、ディスク制御部204とは内部バス207で接続される。CPU 201は、内部バス207を用いて、それらの制御部と、制御情報及びデータのやり取りを行う。

【0042】ポート制御部206は、ファイバチャネル105を介してスイッチ102のポート制御部112と接続されており、そのポート制御部102とコマンド、データ等のやり取りを行う。

【0043】ディスクドライブ205は、ディスク制御部204に接続されオペレーティングシステム208、ポート制御部206等のハードウェアを制御するデバイスドライバ209、アプリケーションプログラム210、及びこれらのプログラムの実行時に必要なデータを格納する。

【0044】メモリ203は、メモリ制御部202と接続される。オペレーティングシステム208、ディスク制御部204、及びデバイスドライバ209、アプリケーションプログラム210等がCPU210により実行される時、ディスクドライブ205から読み出され、メモリ203上に格納される。

【0045】(2) データ移行処理のフロー
次に、図9を参照して、本発明のデータ移行処理を説明する。なお、ホスト101と移行元ディスク装置103とは、既にスイッチ102に接続されているものとする。

【0046】(A) 移行先ディスク装置をスイッチに接続 (901)

まず、オペレータは、制御装置150を起動し、移行元デ

12

ィスク装置103及び移行先ディスク装置104が接続されているスイッチ102のポート番号を入力する。

【0047】次に、オペレータは、移行先ディスク装置104をスイッチ102に接続する。この時点では、移行先ディスク装置104のポートIDには、既に使用されているポートID以外のポートIDであればどのようなポートIDを付けても良い。スイッチ102の各ポートにデフォルトでポートIDが割り当てられているときは、移行先ディスク装置104が接続されるF_Portに与えられているポートIDを

10 与えればよい。本実施例では、上述のように、移行先ディスク装置104には、論理ポートID_2が与えられる。

【0048】移行先ディスク装置104に割り当てられた論理ポートIDは、ホスト101が使用できるディスク装置の論理ポートID_1とは異なるので、この時点では、ホスト101は、移行先ディスク装置104にアクセスすることはできない。

【0049】(B) 移行元ディスク装置の構成情報を移行先ディスク装置に入力 (902)

移行先ディスク装置104の接続後、オペレータは、管理装置150から上述のコマンドにより、移行先ディスク装置104内の移行プログラム142を起動させる。起動された移行プログラム142は、まず、移行元ディスク装置103から構成情報127を取得する(902)。

【0050】図7に示すように、構成情報127は、移行元ポートID 702、World wide Name 703、ファイバ構成情報704、SCSI構成情報705を含む。

【0051】ファイバ構成情報704は、PLOGIペイロード706とPRLIペイロード707とを含む。PRLIペイロード707は、ファイバチャネルのN_Port Login (PLOGI) の際に

30 やり取りされる共通サービス・パラメータ708、ポート名709、ノード名710、及びクラス・サービス・パラメータ711を含む。共通サービス・パラメータ708には、ファイバチャネルのバージョン情報や該ファイバチャネル機器がサポートしているアドレス指定方法等の機能、及び通信方法等が指定されている。クラス・サービス・パラメータ711は、クラスのサポート情報、X_IDの再割当、ACKの能力等を示す。ファイバチャネルのパラメータに関する記載は、上述のANSIのX3.230 FC-PHに詳しく述べて

40 いる。PRLIペイロード707は、サービス・パラメータ712を含む。

【0052】SCSI構成情報705は、INQUIRYデータ713、ディスコネクト・リコネクト・パラメータ714、センスデータ715、及びモード・セレクト・パラメータ716を含む。INQUIRYデータ713は、SCSI装置のタイプ、ベンダID、プロダクトID等を示す。ディスコネクト・リコネクト・パラメータ714は、接続の条件を示す。センスデータ715は、エラー発生時にディスク装置の状態を調べるためにやり取りされるものである。モード・セレクト・パラメータ716は、SCSI装置の物理的属性、記憶媒体上のデータ形式、エラーリカバリの方法や手順、T/Oプロ

50

(8)

13

セスの処理方法などに関する各種のパラメータの設定や変更を行うものである。INQUIRYデータ713とモード・セレクト・パラメータ716により、移行元ディスク装置103内のボリューム数、各ボリュームのサイズ（ブロック数）を知ることができる。ファイバチャネル105上のSCSIの詳細なプロトコルに関しては、ANSIのX3.269 Fibre ChannelProtocol for SCSIに詳しい。

【0053】構成情報127は上記のように既存のファイバチャネル、SCSI等のプロトコルを用いて取得できる情報以外にも存在し得る。そのような構成情報が存在する場合には、オペレーターが、移行元ディスク装置103から構成情報127を直接読み取り、移行先ディスク装置104に入力する。読み取り、入力は、操作パネル、又はHTTPによるwebアクセスにて取得する。

【0054】移行先ディスク装置104の移行プログラム142は、移行先ディスク装置104への構成情報127の入力、すなわち、移行元構成情報138の入力が終了すると、その旨をスイッチ102に通知する。

【0055】(C) ポート切替処理 (903)

スイッチ102のスイッチ制御プログラム119が、その通知を受けると、ポート切換処理が開始される。

【0056】ポート切換処理の詳細を、図10のフローチャートを用いて説明する。

【0057】最初に図3に示したように、ファブリック301が一つのスイッチ102で構成される場合について説明する。

【0058】まず、スイッチ102のスイッチ制御プログラム119が、移行元ディスク装置103を使用している全ホスト101に、ポート切換を開始する旨を通知する。その通知を受け取ったホスト101のデバイスドライバ209は、移行元ディスク装置103に対するI/Oを、ホスト101のメモリ203上にキューイングする(1001)。

【0059】移行元ディスク装置103へのI/Oが停止されると、デバイスドライバ209は、スイッチ102に対してI/Oの停止完了を通知する。移行元ディスク装置103に対して実行中のI/O処理がある場合は、中止させてもよいが、終了するまで実行させた後、I/Oの停止完了が通知されるのが望ましい。スイッチ102のスイッチ制御プログラム119は、全ホスト101からその通知を受け付けた後、ポート構成テーブル118の移行元ディスク装置103の論理ポートID 601と物理ポートID 602との対応関係、移行先ディスク装置104の論理ポートID 601と物理ポートID 602との対応関係を変更する(1002)。つまり、ポート構成テーブル118は、図6(B)のように書き換える。この様子をコンピュータシステム全体の図で表すと、図20に示す状態から図21に示す状態に変化することとなる。

【0060】以後、スイッチ102のポート制御部112は、フレームの送受信毎にポート構成テーブル118を参照することにより、S_ID 508及びD_ID 507を操作し、ポートの切換処理を行う。フレームの受信時には、受信したフ

特開2001-249853

14

レームのS_ID 508に対応する論理ポートID 601が検索され、そのフレームのS_ID 508は、その検索された論理ポートID 601に対応する物理ポートID 602に変換される。同様に、フレーム送信時は、フレームのD_ID 507と対応する論理ポートID 601が検索され、そのD_ID 507は、その検索された論理ポートID 601に対応する物理ポートID 602に変換される。このとき、フレームに付随しているCRC 504は計算し直される。以上の処理により、移行元ディスク装置103に対するフレームは、すべて移行先ディスク装置104に送信される。また、移行先ディスク装置104から送信されたフレームは、ホスト101からは、移行元ディスク装置103から送信されたよう見える。スイッチ102内のポート切換後、ホスト101のI/Oを再開する。(1003) 次に、図4に示したように、ファブリック401が複数のスイッチ102で構成される場合について説明する。

【0061】まず、マスタとなるスイッチ102（マスタスイッチ）を決める。本実施例では移行先ディスク装置104が直接接続されているスイッチ102が、マスタスイッチとなる。

【0062】マスタスイッチのスイッチ制御プログラム119は、移行元ディスク装置103を使用している全ホスト101、及びファブリック401内のマスタスイッチ以外の全スイッチ102に、ポート切換を開始する旨を通知する。その通知を受け取ったホスト101のデバイスドライバ209が行う処理は、I/Oの停止完了をマスタスイッチに通知することを除いて、ファブリックが一つのスイッチ102で構成されている場合と同じである。

【0063】マスタスイッチのスイッチ制御プログラム301118は、全ホスト101からその通知を受け付けた後、ポート構成テーブル118の移行元ディスク装置103の論理ポートID 601と物理ポートID 602との対応関係と、移行先ディスク装置104の論理ポートID 601と物理ポートID 602との対応関係を変更する共に、その変更をマスタスイッチ以外の全スイッチ102に通知する。マスタスイッチ以外の全スイッチは、その通知に基づき、自身のポート構成テーブル118を変更する。以後の各スイッチの動作は、ファブリックが一つのスイッチ102で構成されている場合のスイッチ102の動作と同じである。ファブリック401内の全スイッチ102のポート切換が行われた後、ホスト101のI/Oを再開する。

【0064】(D) データ移行処理 (904)

ポート切替処理が終了に同期して、データ移行処理が行われる。図11のフローチャートを用いて、そのデータ移行処理を説明する。

【0065】まず、移行プログラム142は、移行元ディスク装置103上の論理ボリューム番号とそのボリュームの大きさに応じたボリュームを移行先ディスク装置104に構築し、使用する変数及び図8で説明した移行ワークテーブル143の初期化を行う(1101)。

50

(9)

特開2001-249853

15

【0066】移行先ディスク装置104上の移行プログラム142は、ホスト101からのI/O要求の有無をチェックする(1102)。

【0067】ホスト101からのI/O要求が無い場合は、移行元ディスク装置103から、スロット単位でデータを移行する。その際、移行プログラム142は、移行ワークテーブル143を用いて、次に移行するデータのアドレスを算出する(1103)。

【0068】移行元ディスク装置103の全論理ボリュームに対して、ボリューム番号801の小さい順からデータ移行が実行され、また、各ボリューム内においては、先頭ブロックからデータ移行が行われる。図8に示した移行用ワークテーブル132は、その先頭行のボリューム番号801の欄に次移行アドレスを示す。次移行アドレスの初期値は、最小のボリューム番号801の先頭ブロックアドレスとなる。また、移行ワークテーブル143は、その先頭行のスロット番号802の欄に次移行スロットを示す。2番目以降の行はボリューム番号801、スロット番号802の昇順となるよう整列される。移行プログラム142は、移行ワークテーブル143から得た次移行アドレスのデータを移行する際、そのスロット番号のステータス803を「移行中」へ変更する(1104)。

【0069】次に移行プログラム142は、ポート制御部125を用いて移行元ディスク装置103へスロットのサイズのリード要求を発行し、該当するデータを取得する(1105)。

【0070】移行プログラム142は、移行元ディスク装置103から得たデータを移行先ディスク装置104の対応するボリュームへディスク制御部124を使用しI/F 140を介してディスクドライブ141へ書き込む(1106)。

【0071】ディスクドライブ141へデータを書き込んだ後、当該ボリュームの全スロットの移行が終了したかチェックし、終了している場合には次のボリュームのデータ移行を実行し、終了していない場合はスロット番号をインクリメントする(1107)。

【0072】全ボリュームに関する移行が終了している場合には移行処理を終了する(1108)。

【0073】次にホスト101からのI/O要求があった場合について説明する。ホスト101からのI/O要求がある場合には、移行プログラム142はリードかライトか調べる(1109)。

【0074】移行プログラム142はそれがリードの場合、移行ワークテーブル143を参照し、要求データが移行済みかどうか調べる(1110)。

【0075】移行プログラム142は、移行済みの場合には、移行先ディスク装置104内のディスクドライブ141から読み出し、ホスト101に返送する(1113, 1112)。

【0076】ホスト101から要求されたデータが移行済みでない場合には、移行プログラム142はポート制御部125を用いて、スイッチ102経由で移行元ディスク装置103

50

16

にリード要求を発行し、当該データを取得する。移行プログラム142は、移行元ディスク装置103から取得したデータをホスト101へ返送すると共に、移行先ディスク装置104の対応するボリュームへそのデータを書き込む(1111, 1112)。

【0077】ディスクドライブ141へ当該データを書き込んだ後、該データを移行ワークテーブル143に登録し、ステータス803を移行済みとし、該要求に対する処理を終える(1108)。

10 【0078】次に、ホスト101からのライト要求があつた場合について説明する。移行プログラム142は、ホスト101からライト要求に関するデータを受け取る(1114)。

【0079】次に、移行プログラム142は、ポート制御部125を用いて、スイッチ102経由で、移行元ディスク装置103へホスト101からの要求データを含むようスロットサイズ単位のリード要求を発行しデータを取得する。移行プログラム142は移行元から得られたデータにホスト101から受け取ったデータを上書きし新しいスロットデータを作成した後、移行先ディスク装置104内のホスト101が要求したボリュームに対して書き込みを行う(1115)。

20 【0080】書き込みが終了したら、移行プログラム142は移行ワークテーブル143へ該スロットを登録し、ステータス803を「移行済み」とし、ホスト101からのライト要求に関する処理を終える(1107)。

【0081】(E) 移行元ディスク装置の切り離し(905)

データ移行処理の終了後、オペレータは移行元ディスク装置103をスイッチから切り離し、移行処理を終了する(905)。

30 【0082】以上、実施例1のシステム構成及びデータ移行処理のフローを説明した。本実施例によれば、ホストからのI/O要求が停止されるのは、移行元ディスク装置103の構成情報を移行先ディスク装置104に入力されている間、及びポートの切替処理が行われている間だけである。従って、本実施例によれば、ホストと移行元ディスク装置との間の接続を切り離し、移行先ディスク装置をホスト及び移行元ディスクに接続するという作業は必要ないので、ホストからのI/O要求が停止される時間は、米国特許第5,680,640号に開示されるデータ移行システム及びその方法においてホストからのI/O要求が停止される時間に比べて短い。そのため、データ移行に伴うコストをより抑えることができる。さらに、本実施例によれば、データ移行のための専用チャネルを設ける必要ないので、ハードウェアのコストも、より抑えることができる。

【0083】【実施例2】本実施例は、実施例1で説明したシステム構成において、さらに、移行元ディスク装置と移行先ディスク装置とを、インターフェース(例え

(10)

17

ばファイバチャネル又はSCSI)で直接接続する点で、実施例1と異なる。本実施例では、そのインターフェースを用いてデータ移行を行うので、実施例1に比べて高速にデータ移行を行うことができる。本実施例では、データ移行用のインターフェースを設ける必要がある。しかし、ホストと移行元ディスク装置との間の接続を切り離し、移行先ディスク装置をホスト及び移行元ディスクに接続するという作業は必要なく、単に、移行先ディスク装置をスイッチに接続し、移行元ディスク装置と移行先ディスク装置とをインターフェースで接続するだけなので、上述の米国特許第5,680,640号に開示されるデータ移行システム及びその方法に比べて、移行先ディスク装置を追加する作業が短縮される。

【0084】以下、図12を用いて、本実施例を説明する。

【0085】図12は、本発明のデータ移行システム及びその方法が適用されるコンピュータシステムの第2の実施例を説明するための図である。上述のように、そのコンピュータシステムの各構成要素は、図1の各構成要素と同様であるので、説明は省略する。

【0086】以下、本実施例におけるデータ移行処理フローを説明する。データ移行処理フローは、図9で説明した実施例1のものと同様である。但し、本実施例では移行元ディスク装置1203と移行先ディスク装置1204とが直接インターフェースで接続されており、そのインターフェースを用いてデータ移行を行うので、移行元構成情報1238により、移行元ディスク装置1203内の論理アドレスと物理アドレスの対応関係を求める必要がある。以下、その異なる点のみを説明する。

【0087】図11のステップ1105、1113、及び1115では、スイッチ102を介して移行元ディスク装置からデータを読み込んだ。しかし、本実施例では、移行プログラム1231は、移行元構成情報1238に含まれる情報を用いて、移行元ディスク装置1203から読み出すべきデータの論理アドレスに対応する物理アドレスを算出する。そして、移行プログラム1231は、I/F1240を用いて、算出した物理アドレスに対応する移行元ディスク装置1203のディスクドライブ1230に対して、直接リード要求を発行し、当該データを取得する。このとき、当該データが移行元ディスク装置1203の複数のディスクドライブ1230に分散されている場合、又は離散した物理アドレスに格納されている場合は、複数のリード要求を発行する。

【0088】本実施例では、移行元ディスク装置1203と移行先ディスク装置1204とを直接インターフェースで接続し、スイッチ1202を介さず、そのインターフェースを用いてデータ移行を行うため、実施例1に比べて高速にデータ移行を行うことができる。また、本実施例のデータ移行はスイッチ1202を介さないので、スイッチのリソースの消費を抑えることができる。本実施例では、スイッチ1202に接続されているのは、ホストが2台、ディスク

特開2001-249853

18

装置が2台であるが、SAN環境下では、多くのホスト及びディスク装置が接続されるので、そのリソースを他のホストと他のディスク装置間のデータ転送に割り当てることが可能となる。従って、本実施例によれば、実施例1に比べて、データ移行に伴うスイッチの性能低下を抑制することができる。

【0089】【実施例3】本実施例では、移行プログラムをスイッチで実行する点で実施例1と異なる。本実施例によれば、ディスク装置に移行プログラムを持たせる必要がないので、そのようなプログラムを持たないディスク装置でも移行元ディスク装置からのデータ移行が可能となる。

【0090】図13を用いて、本実施例のコンピュータシステムの構成を説明する。但し、図1の構成要素と同様のものについては説明を省略する。

【0091】本実施例では、スイッチ1302内の不揮発メモリ1317に、スイッチ1302の制御に必要なスイッチ制御プログラム1319、移行処理を行う移行プログラム1342と、それらのプログラムの実行時に必要なポート構成テーブル1318、移行元構成情報1327、及び移行ワークテーブル1343等のデータとを格納する。移行先ディスク装置1304内のメモリ1337にはディスク装置制御プログラム1349が格納されるが、実施例1のように移行元構成情報や移行ワークテーブル、移行プログラムは格納されない。

【0092】図14は、ポートの構成を示すポート構成テーブル1318の構成例である。ポート構成テーブル1318は、仮想的なポートIDを示す論理ポートID 1401、物理的なポートIDを示す物理ポートID 1402、及びポートの状態を示すステータス1403とを含む。

【0093】次に、図15のフローチャートを用いて、本実施例のデータ移行処理を説明する。

【0094】オペレーターは、移行先ディスク装置1304をスイッチ1302に接続する(1501)。このステップは、図9のステップ901と同様である。

【0095】次に、スイッチ1302の移行プログラム1342は、移行元ディスク装置1303から構成情報1327を読み取る(1502)。構成情報1327は図7に示すものと同様である。

【0096】移行元ディスク装置1303の構成情報1327がスイッチ1302に入力された後、スイッチ1302のスイッチ制御プログラム1319はポート切換処理を開始する(1503)。

【0097】本実施例のポート切替処理においても、実施例1と同様に、ホストのI/Oを止めた後、ポート構成テーブル1318の移行元ディスク装置1303の論理ポートID 1401と物理ポートID 1402との対応関係と、移行先ディスク装置1304の論理ポートID 1401と物理ポートID 1402との対応関係を変更する。しかし、本実施例のポート切替処理では、次の点で実施例1のそれと異なる。本実施例では、スイッチ1302のスイッチ制御プログラム1319は、

(11)

19

ポート構成テーブル1318における移行元ディスク装置1303のポートIDのステータス1403を「移行中」とする。ポート構成テーブル1318のステータス1403が「移行中」のポートIDに対するフレームは、対応するポートには送信されず、スイッチ1302の移行プログラム1331に渡される。つまり、移行元ディスク装置1303へのアクセスはスイッチ1302の移行プログラム1331が行う。

【0098】ファブリックが複数のスイッチ1302により構成されている場合には、実施例1と同様のやり方でホストのI/Oを止めた後、マスタスイッチのスイッチ制御プログラム1319は、ポート構成テーブル1318における移行元ディスク装置1303のポートIDのステータス1403を「移行中」とする。ポート構成テーブル1318のステータス1403が「移行中」のポートIDに対するフレームは、対応するポートには送信されず、スイッチ1302の移行プログラム1331に渡される。ファブリックが複数のスイッチ1302で構成されている場合には、少なくともマスタスイッチとなるスイッチ1302に移行元構成情報1327、移行ワークテーブル1343、及び移行プログラム1342が搭載されれば足りる。

【0099】ポート切替処理が終了した後、データ移行処理を行う。まず、スイッチ1302の移行プログラム1342は、移行元ディスク装置1303上に構成されている論理ボリューム番号と該ボリュームの大きさに従ったボリュームを構築し、使用する変数や図8に示す移行ワークテーブル1332の初期化を行う。以後の処理は、移行先ディスク装置の移行プログラムではなく、スイッチ1302の移行プログラム1302が行うという点を除き、実施例1で説明した処理と同様である。

【0100】データ移行処理が終了した後、スイッチ1302のスイッチ制御プログラム1319は、ポート構成テーブル1318の移行先ディスク装置1304のステータスを「通常」に変更する(1505)。オペレータは移行元ディスク装置1303をスイッチ1302から切り離し、移行処理を終了する(1506)。

【0101】[実施例4] 本実施例の特徴は、スイッチに接続されるパーソナルコンピュータ、ワークステーション等で構成されるマイグレータ上で移行プログラムが実行される点である。本実施例によれば、実施例3と同様に、移行先ディスク装置に移行プログラムを持たせる必要がないので、そのようなプログラムを持たないディスク装置でも移行元ディスク装置からのデータ移行が可能となる。さらに、本実施例では、実施例3のように、スイッチ上で移行プログラムを実行させないので、スイッチの負荷を減らすことが可能となる。

【0102】図16を用いて、本実施例のコンピュータシステムの構成を説明する。但し、図1の構成要素と同様のものについては説明を省略する。

【0103】本実施例では、後述するようにマイグレータ1606内に、移行処理を行う移行プログラムと、移行元

特開2001-249853

20

構成情報、及び移行ワークテーブルを格納する点に特徴がある。移行先ディスク装置1604内のメモリ1703にはディスク装置制御プログラム1339が格納されるが、実施例1のように移行元構成情報や移行ワークテーブル、移行プログラムは格納されない。また、実施例3と異なり、本実施例では、スイッチ1602に、移行プログラムと、移行元構成情報、及び移行ワークテーブルは格納されない。

【0104】図17は、マイグレータ1606の構成例を示す。マイグレータ1606はCPU 1701、メモリ制御部1702、ポート制御部1706、ディスク制御部1704、ディスクドライブ1705、メモリ1703を含む。CPU 1701と、メモリ制御部1702、ポート制御部1706、及びディスク制御部1704とは内部バス1707で接続され、CPU 1701は各制御部との間で制御情報、データのやり取りを行う。ポート制御部1706は、ファイバチャネル1605によりスイッチ1602のポート制御部1612と接続され、スイッチ1602とコマンド、データのやり取りを行う。マイグレータ1606のポートはNポートと呼び、スイッチ1602側のポートをFポートと呼ぶ。ディスクドライブ1705はディスク制御部1704と接続され、CPU 1701からのリード・ライトの指示を受け付ける。ディスクドライブ1705にはマイグレータ1606に必要なプログラムが格納されており、移行プログラム1731及び、該プログラムの実行時に必要な移行ワークテーブル1732等が格納する。メモリ1703はメモリ制御部1702と接続され、移行プログラム1731の実行時に当該プログラムがドライブ1705からメモリ1703上に読み出される。

【0105】次に、本実施例のデータ移行処理のフローを説明する。オペレータは移行先ディスク装置1604をデータ移行用スイッチ1602に接続する。この時には、移行先ディスク装置1604のポートIDはスイッチ1602のF_Portに与えられたポートIDと同じ値を取り、移行先ディスク装置1604はホスト1601からは認識されない。このステップは、図9で説明したステップ901と同様である。

【0106】次に、マイグレータ1606の移行プログラム1631は、移行元ディスク装置1603から構成情報1727を読み取る。構成情報1727は図7に示すものと同様である。

【0107】移行元ディスク装置1603の構成情報1727がマイグレータに入力された後、ポートの切換処理を行う。その処理は、実施例1で説明したものと同様に、ホストからのI/Oを停止させ、ポート構成テーブル1618の移行元ディスク装置1603の物理ポート1604と移行先ディスク装置1604の物理ポートID1402とを入れ替える。スイッチ1602の制御プログラムはポート構成テーブル1618の移行元ディスク装置1603のポートIDのステータスを「移行中」とする。ポート構成テーブルのステータスが「移行中」の場合はフレームポートに送信されずにマイグレータ1606の移行プログラム1731に渡される。移行元ディスク装置1603へのアクセスはマイグレータ1606の移行プログラム1731が行う。これらの処理の後、ホスト1601の

50

(12)

特開2001-249853

21

I/Oを再開し、マイグレータ1606の移行プログラム1731は移行処理を開始する。

【0108】ポート切替処理終了後、データ移行処理が行われる。最初に、マイグレータ1606の移行プログラム1731は移行元ディスク装置1603上に構成されている論理ボリューム番号と該ボリュームの大きさに従ったボリュームを構築し、使用する変数や、移行ワークテーブル1732の初期化を行う。以後の処理は、移行先ディスク装置の移行プログラムではなく、マイグレータ1606の移行プログラム1731が行うという点を除き、実施例1で説明した処理と同様である。

【0109】データ移行処理が終了した後、実施例3と同様に、スイッチ1602のスイッチ制御プログラム1619は、ポート構成テーブル1318の移行先ディスク装置1304のステータスを「通常」に変更する。オペレータは移行元ディスク装置1603をスイッチ1602から切り離し、移行処理を終了する。

【0110】なお、本実施例ではスイッチと、マイグレータとの間をファイバチャネルで接続したが、ベンダユニークなバス等で接続しても本実施例による効果をえられることは、本発明の技術分野に属する者であれば容易にわかることである。

【0111】【実施例5】本実施例では、ホストがスイッチのポート切換を実施する点に特徴がある。本実施例によれば、実施例1乃至実施例4のように、スイッチ上にポート構成テーブルを搭載する必要がない。従って、このような機能を有さないスイッチを用いてSANを構成した場合でも、データ移行を行うことが可能となる。

【0112】図18を用いて、本実施例のコンピュータシステムの構成例を説明する。但し、図1の構成要素と同様のものについては説明を省略する。後述するように、本実施例では、実施例1と異なり、ホスト1801上にポート構成テーブルが搭載され、スイッチ1802にはポート構成テーブルは搭載されない。

【0113】図19は、ホスト1801の構成例を示す。ホスト1801はCPU 1901、メモリ制御部1902、ポート制御部1906、ディスク制御部1904、ドライブ1905、及びメモリ1903から構成され、CPU 1901と、メモリ制御部1902、ポート制御部1906、及びディスク制御部1904との間は内部バス1907で接続されている。CPU 1901は各制御部との間で制御情報、データのやり取りを行う。ポート制御部1906はファイバチャネル1905によりスイッチ1902のポート制御部1912と接続されており、スイッチ1802との間でコマンド、データのやり取りを行う。ホスト1801のポートをNポートと呼べ、スイッチ1902側のポートをFポートと呼ぶ。ドライブ1905はディスク制御部1904と接続されており、CPU 1901からのリード・ライトの指示を受け付ける。ドライブ1905にはホスト1801に必要なプログラムが格納されており、オペレーティングシステム1908、デバイスドライバ1909、アプリケーションプログラム1910及

10

22

び、ポート構成テーブル1918が格納される。メモリ1903はメモリ制御部1902と接続されており、CPU 1901からのリード・ライトの要求を受け付ける。オペレーティングシステム1908、ディスク制御部1904、ハードウェアを制御するデバイスドライバ1909、アプリケーション1910等のプログラムは、各プログラムの実行時にディスクドライバ1905からメモリ1903上に読み出される。

【0114】次に、本実施例にデータ移行処理のフロー説明する。そのフローは、図9に示した実施例1のデータ移行処理のフローと同様である。但し、本実施例では、ポート切替処理が実施例1と異なる。ここでは、ポート切替処理のみ説明する。

【0115】本実施例のポート切替処理は、ファブリック401が単体のスイッチで構成されている場合でも複数のスイッチで構成されている場合も同様の動作を行う。最初にオペレータの指示により、ホストのデバイスドライバ1909は、移行元ディスク装置1803を使用している全ホスト1801に対して、ポート切換を開始する旨を通知する。その通知を受け取ったホスト1801のデバイスドライバ1909は移行元ディスク装置1803に対するI/Oをいったん蓄え、ホスト1801のメモリ203上にキューイングする。一方、移行元ディスク装置1803に対して実行中のI/O処理は、終了するまで実行される。移行元ディスク装置1803へのI/Oが停止されると、デバイスドライバ1909は、スイッチ1802に対してI/Oの停止完了を通知する。ポートを切り替える旨を通知したホストのデバイスドライバ1909は、全ホスト1801からのI/O停止完了の通知を受け付けた後、ポート構成テーブル1918の移行元ディスク装置1803の物理ポートID 602と移行先ディスク装置1804の物理ポートID 602を入れ替えるように、全ホストに通知する。以後、ホストのデバイスドライバはフレームの送受信毎にポート構成テーブル1918を参照し、S_ID 508、D_ID 507を操作しポートの切換処理を行う。この処理は、実施例1で説明した処理と同様である。

【0116】以上、本発明の実施例1乃至5を説明したが、これらの実施例を適宜組み合わせた形態も考えられる。例えば、実施例3乃至5において、実施例2と同様に移行元ディスク装置1303と移行先ディスク装置1304とを移行用のディスクインターフェースで接続し、これを介してデータ移行を行うことが考えられる。

40
40

【0117】

【発明の効果】以上説明したように、本発明によれば、SAN環境に好適なデータ移行方法及びその装置を提供することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施例におけるコンピュータシステムの構成を説明するための図である。

【図2】ホストの構成を説明するための図である。

【図3】一つのスイッチで構成されるファブリックを説明するための図である。

50

(13)

特開2001-249853

23

【図4】複数のスイッチで構成されるファブリックを説明するための図である。

【図5】ファイバチャネルのフレームを説明するための図である。

【図6】本発明のポート構成テーブルの構成例を示す。

【図7】本発明の移行元構成情報の構成例を示す。

【図8】本発明の移行ワークテーブルの構成例を示す。

【図9】本発明のデータ移行処理のフローチャートである。

【図10】本発明のポート切換処理のフローチャートである。

【図11】本発明のデータ移行処理のフローチャートである。

【図12】本発明の第2の実施例におけるコンピュータシステムの構成を説明するための図である。

【図13】本発明の第3の実施例におけるコンピュータシステムの構成を説明するための図である。

【図14】本発明の第3の実施例におけるポート構成テーブルの構成例を示す。

【図15】本発明の第3の実施例における移行処理のフローチャートを示す。

10

20

*

24

* 【図16】本発明の第4の実施例におけるコンピュータシステムの構成を説明するための図である。

【図17】本発明のマイグレータの構成例を示す。

【図18】本発明の第5の実施例におけるコンピュータシステムの構成を説明するための図である。

【図19】本発明の第5の実施例におけるホストの構成を説明するための図である。

【図20】本発明のポート切替処理前の、各ポートに割り当たれている物理ポートIDと論理ポートIDとの対応を説明するための図である。

【図21】本発明のポート切替処理後の、各ポートに割り当たれている物理ポートIDと論理ポートIDとの対応を説明するための図である。

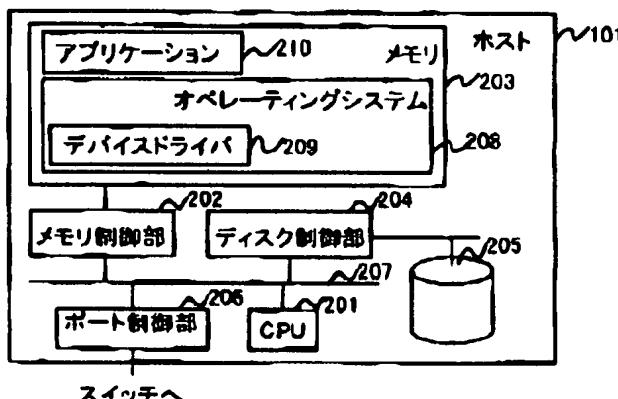
【図22】本発明の制御装置をLANに接続した場合の構成例を示す。

【符号の説明】

101, 1201, 1301, 1601, 1801…ホストコンピュータ, 1
02, 1202, 1302, 1602, 1802…スイッチ, 103, 1203, 13
303, 1603, 1803…移行元ディスク装置, 104, 1204, 13
04, 1604, 1804…移行先ディスク装置, 105, 1205, 130
5, 1605, 1805…ファイバチャネル。

【図2】

図2



スイッチへ

【図8】

図8

ボリューム番号	スロット番号	ステータス
1	10	移行中
1	25	移行済み
1	28	移行済み
2	15	移行済み
2	18	移行済み
:	:	:

【図6】

図6

論理ポートID	物理ポートID
0	0
1	1
2	2
3	3
4	4
N-1	N-1

(A) ポート切替初期値

論理ポートID	物理ポートID
0	0
1	2
2	1
3	3
4	4
N-1	N-1

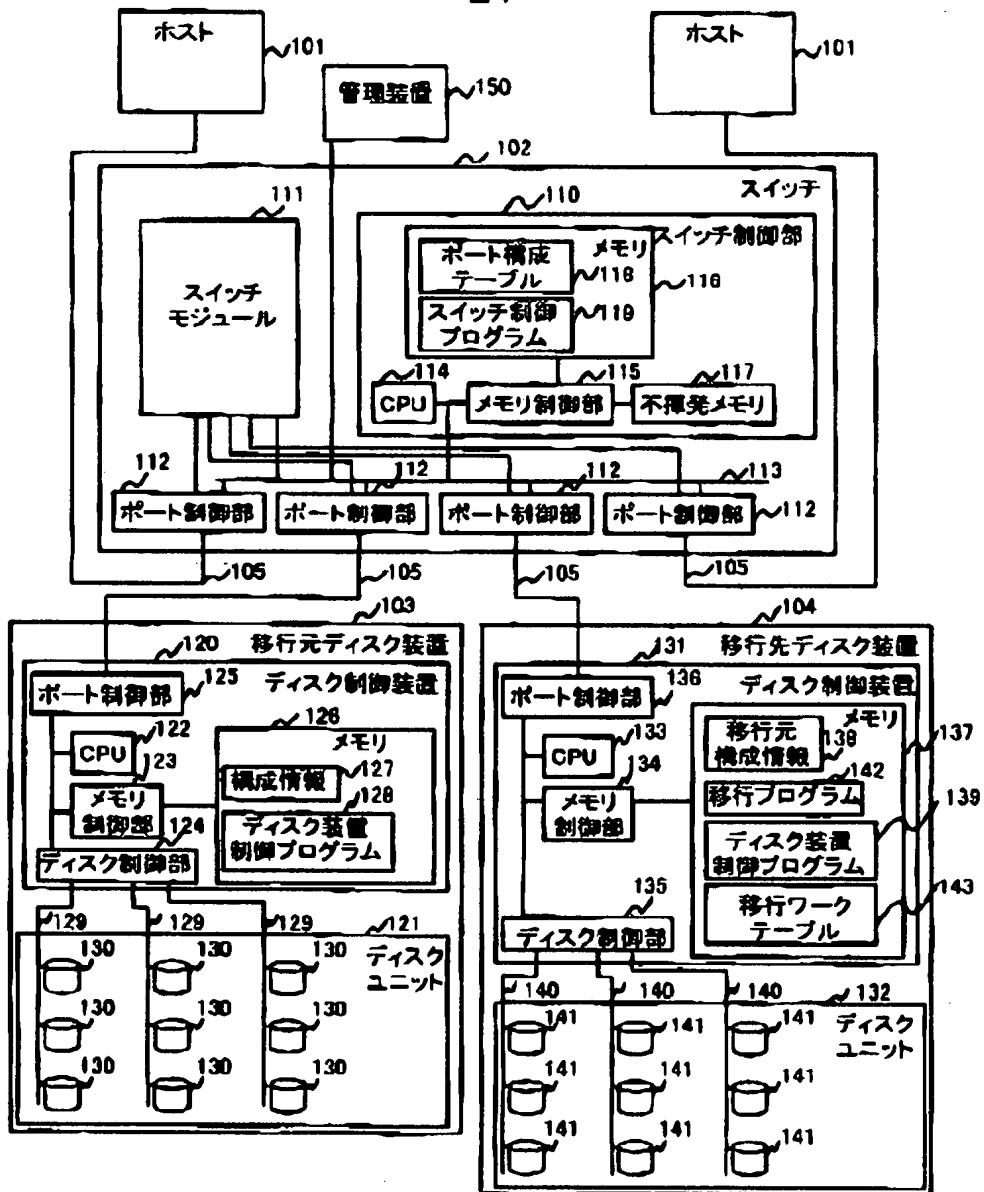
(B) ポート切替処理後

(14)

特開2001-249853

【図1】

図1

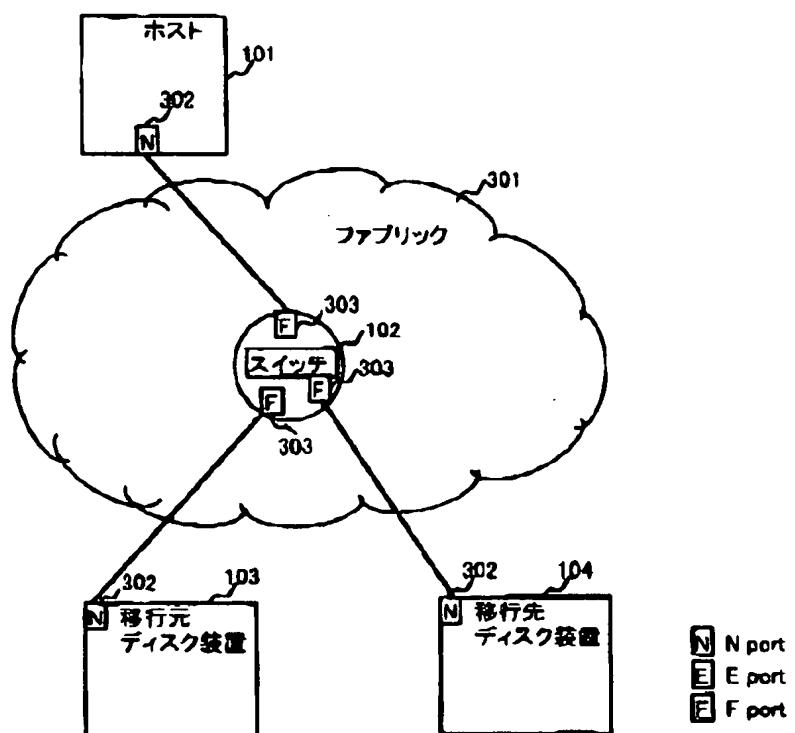


(15)

特開2001-249853

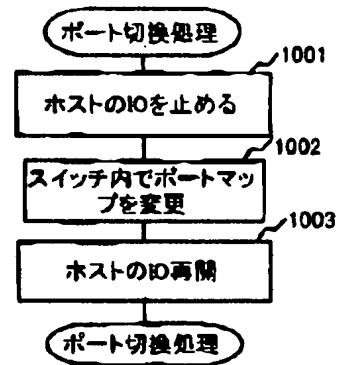
【図3】

図3



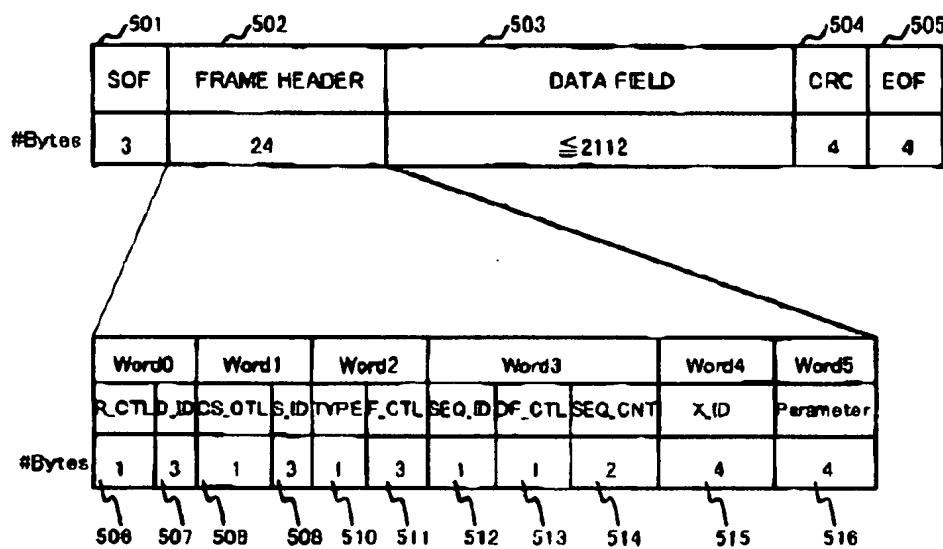
【図10】

図10



【図5】

図5

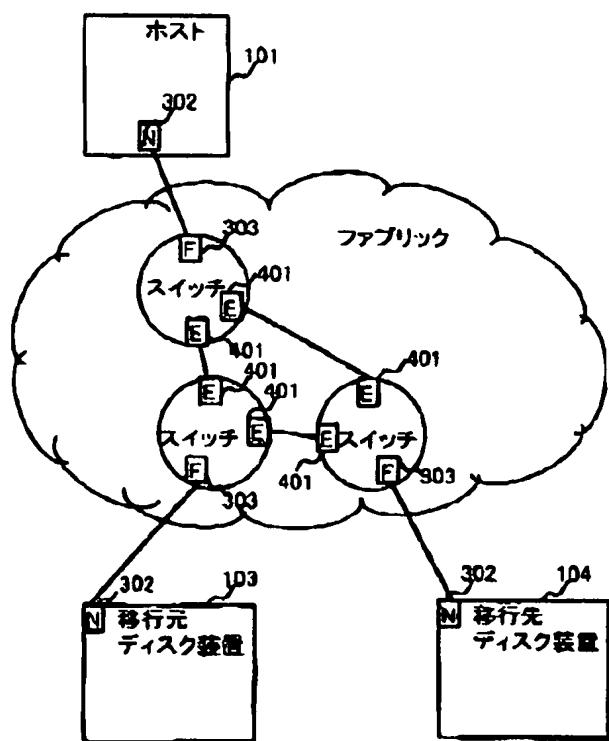


(16)

特開2001-249853

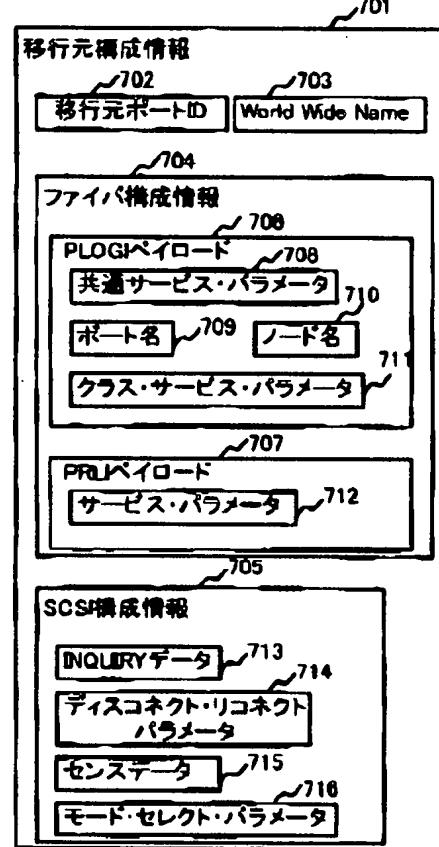
【図4】

図4



【図7】

図7



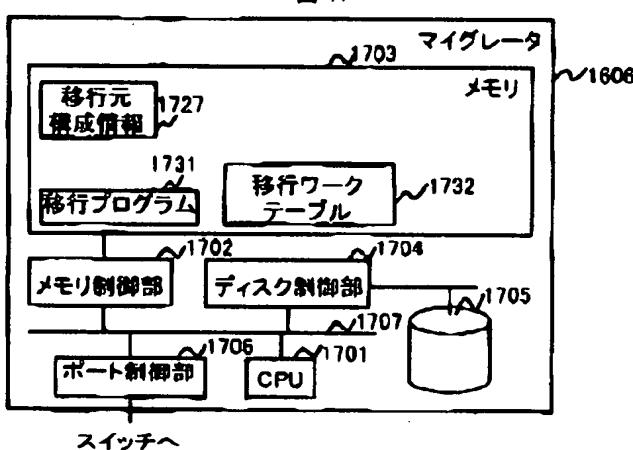
【図14】

図14

物理ポートID	物理ポートID	ステータス
0	0	通常
1	2	移行中
2	1	通常
3	3	通常
⋮	⋮	⋮
N-1	N-1	通常

【図17】

図17

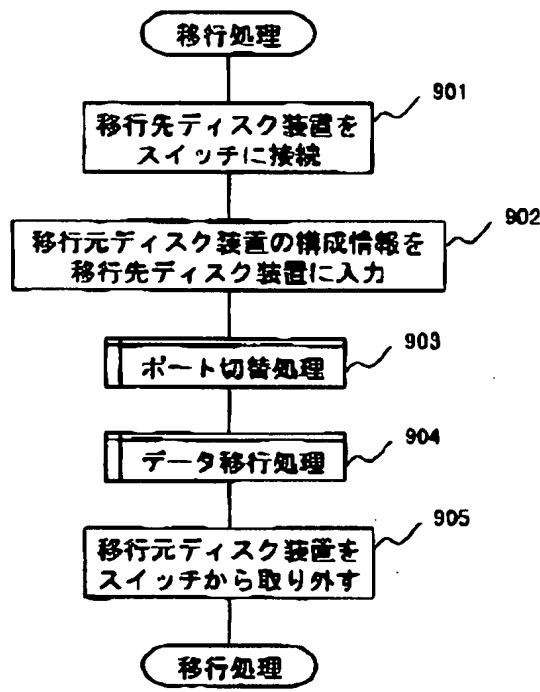


(17)

特開2001-249853

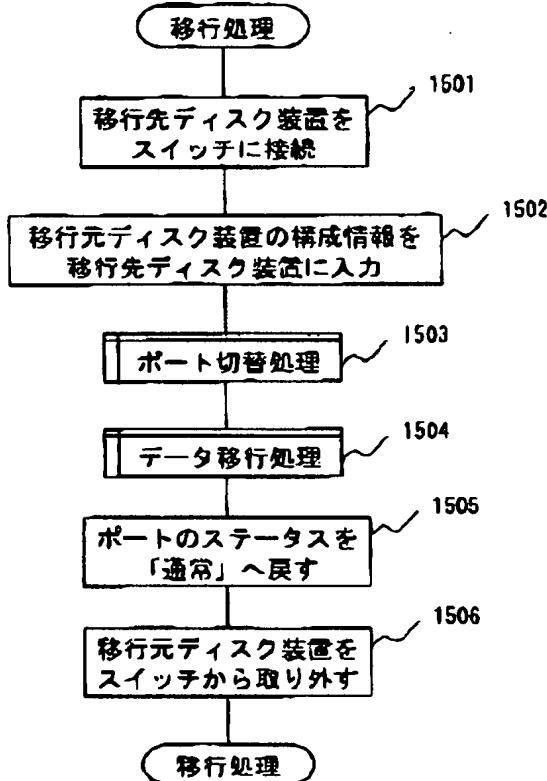
【図9】

図9



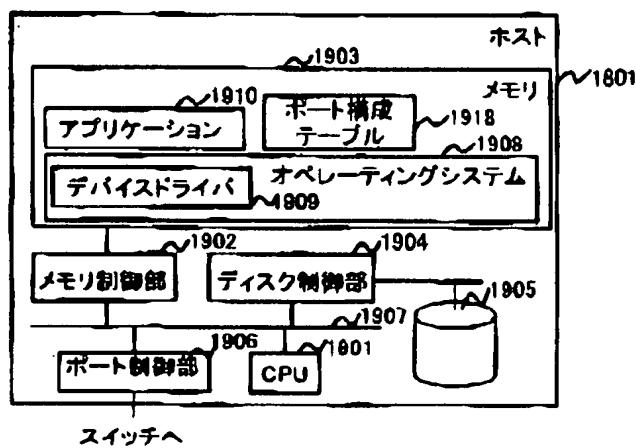
【図15】

図15



【図19】

図19

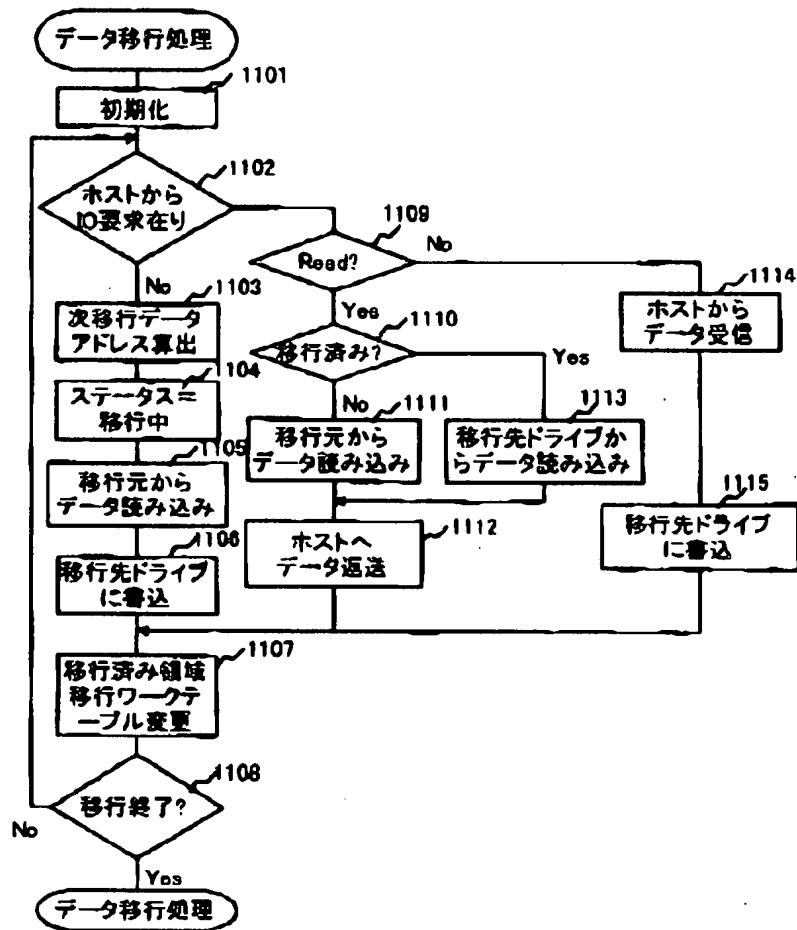


(18)

特開2001-249853

〔図11〕

図11

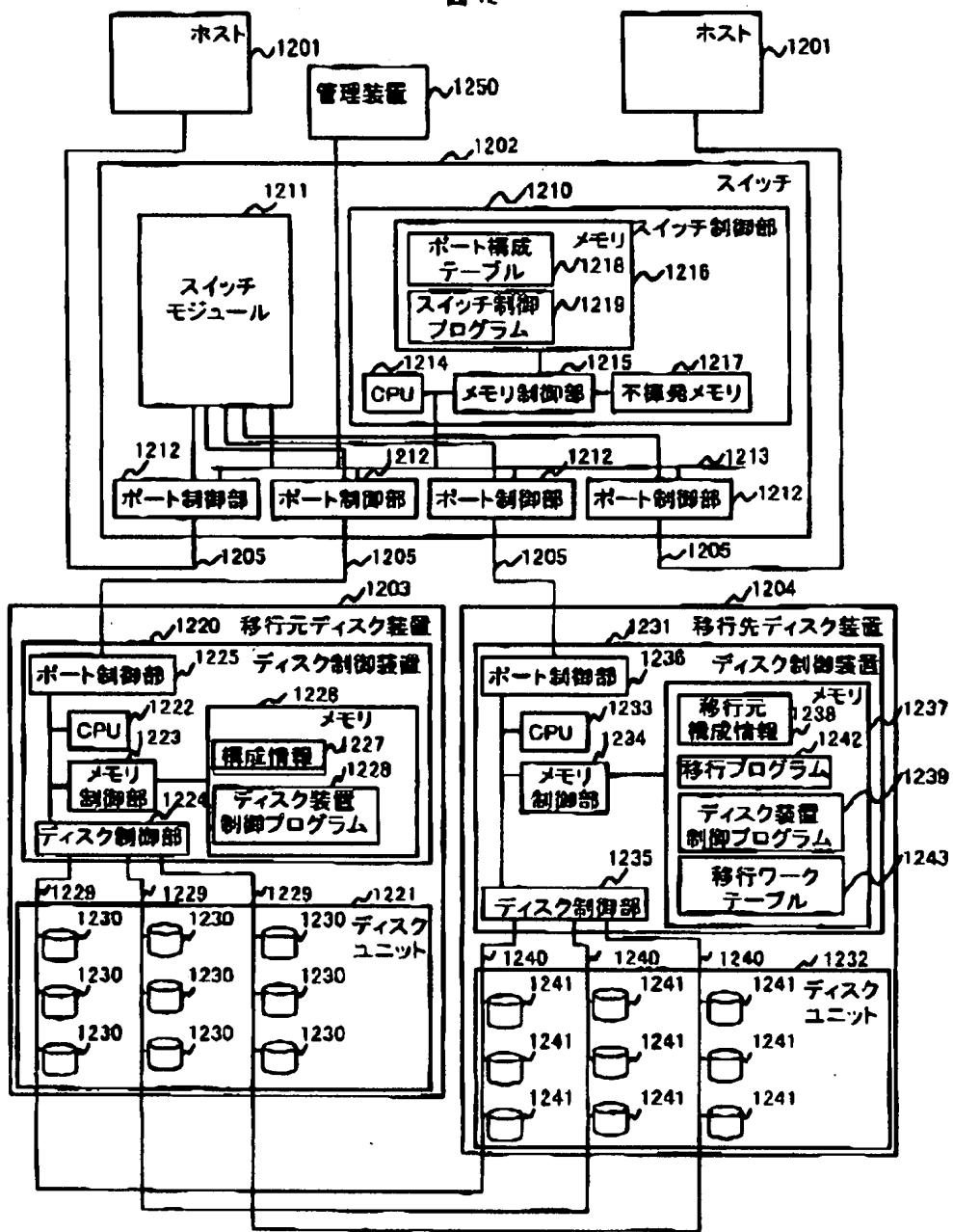


(19)

特開2001-249853

[図12]

図12

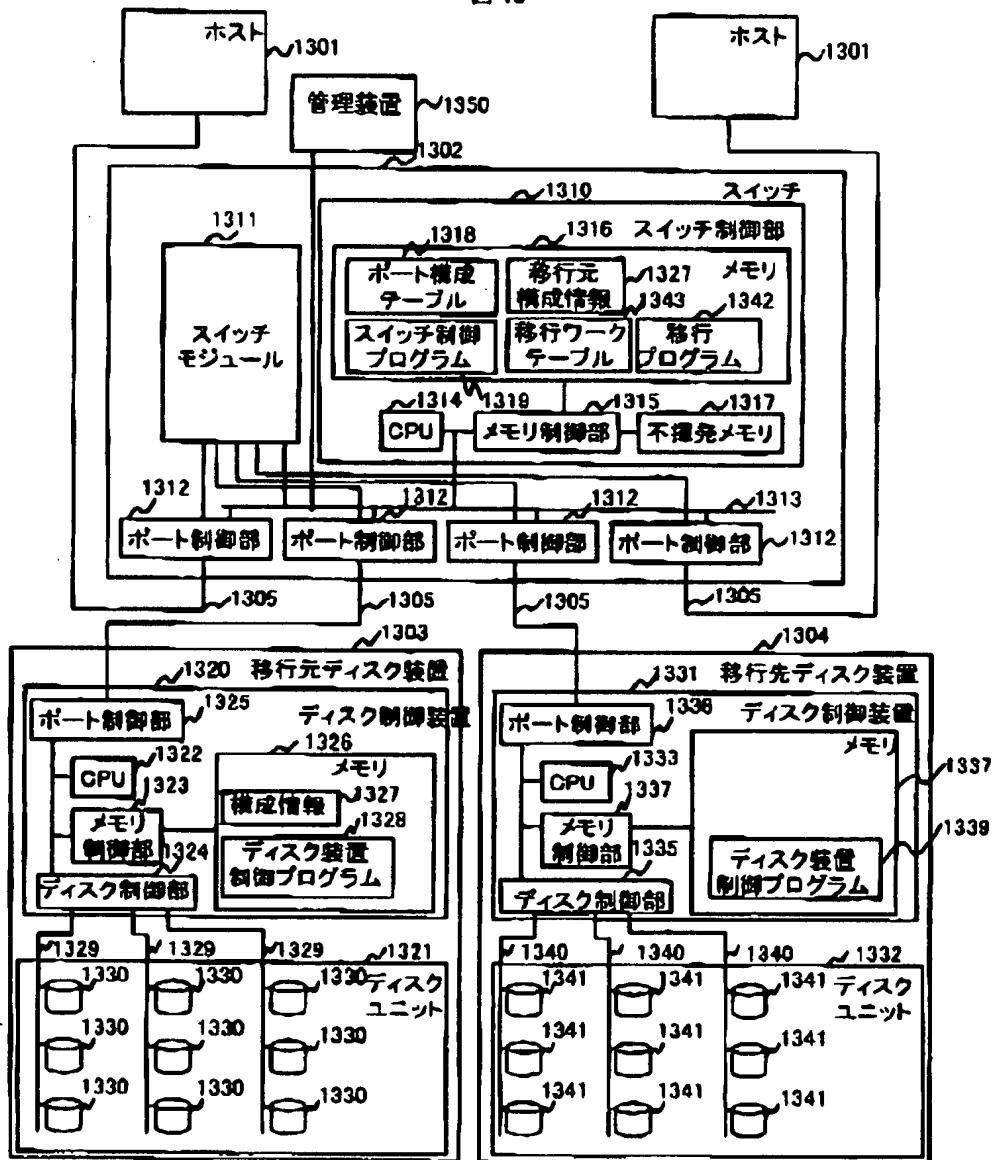


(20)

特開2001-249853

[図13]

図13

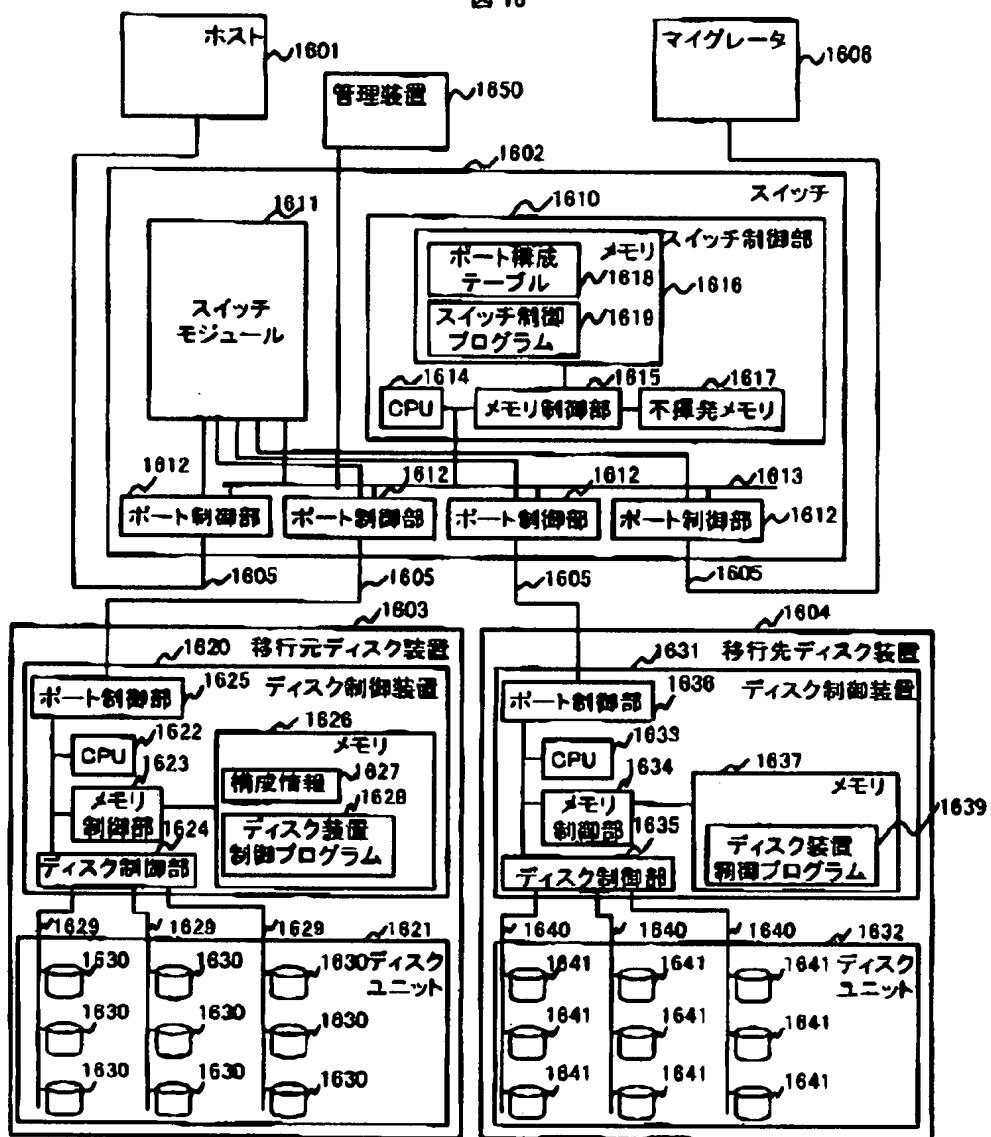


(21)

特開2001-249853

【図16】

図16

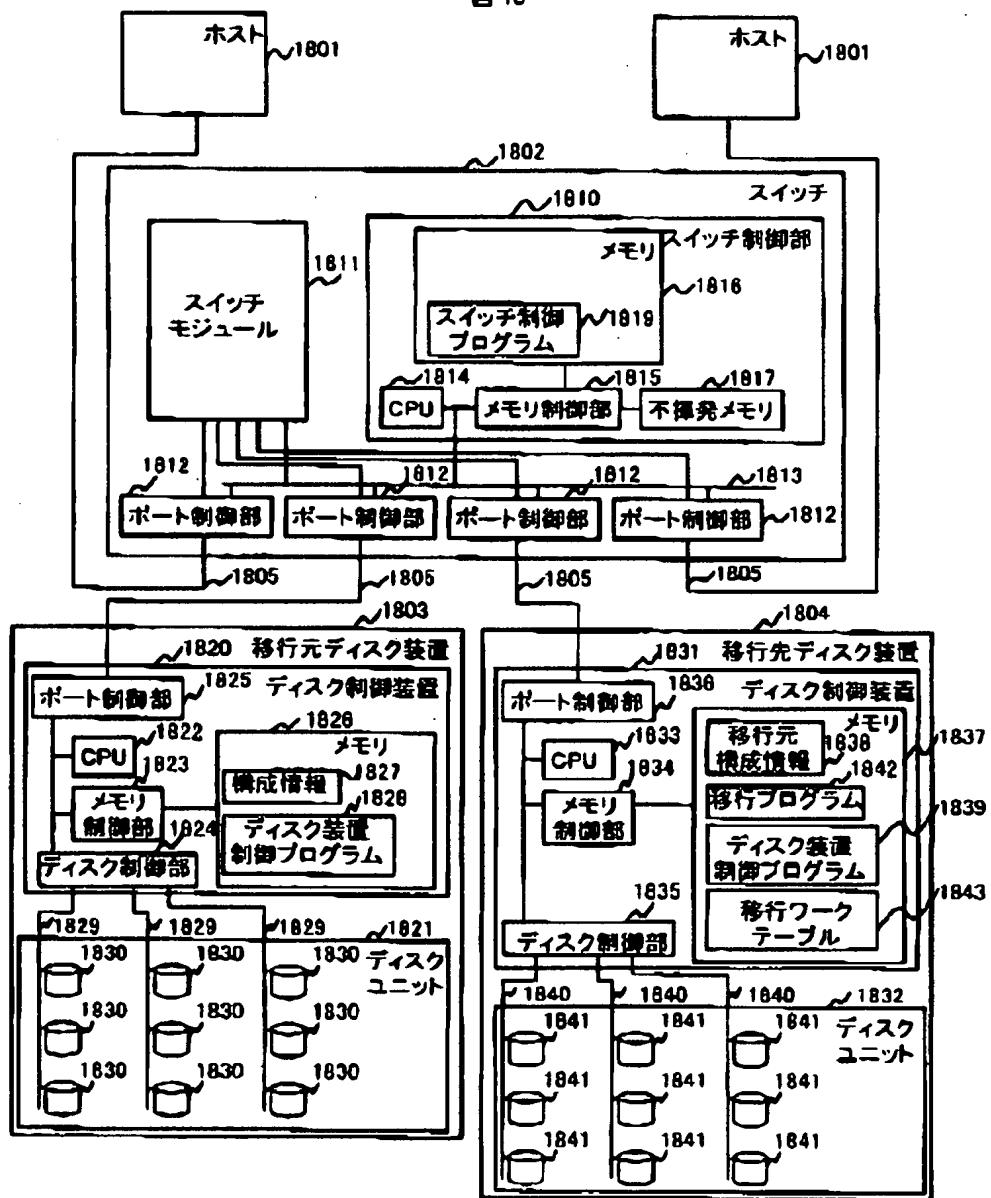


(22)

特開2001-249853

【図18】

図18

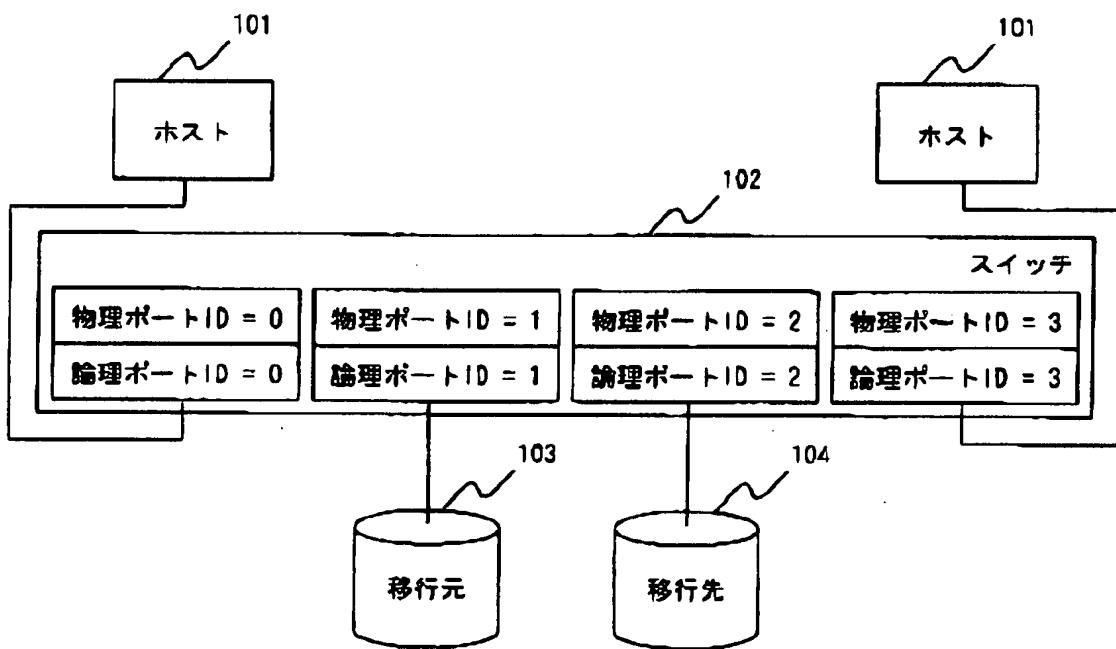


(23)

特開2001-249853

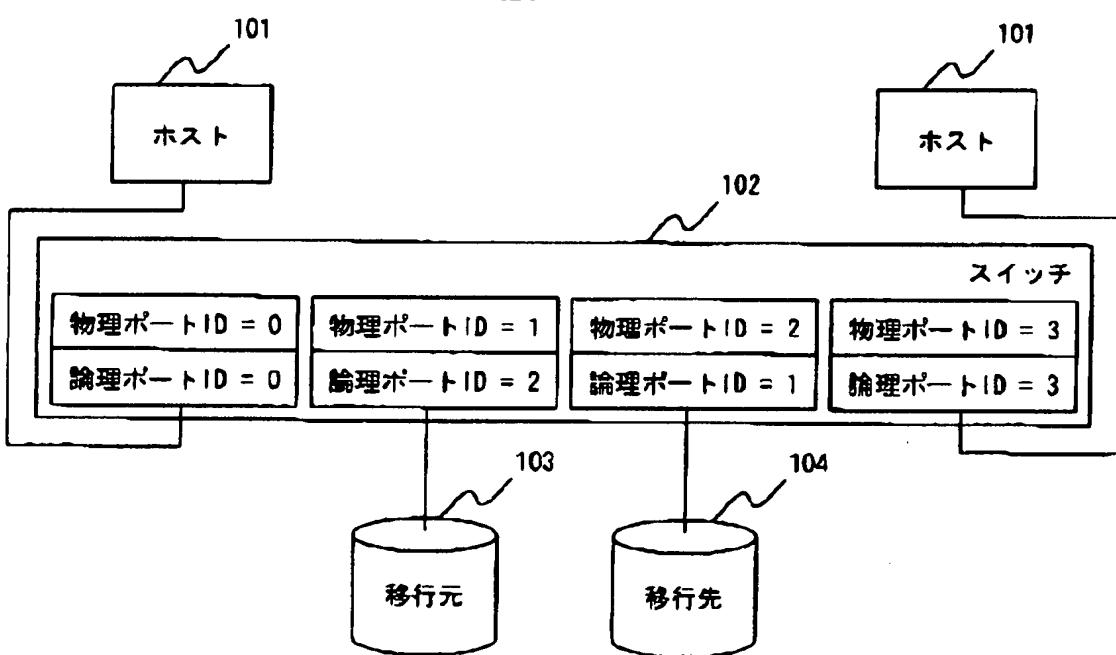
【図20】

図20



【図21】

図21

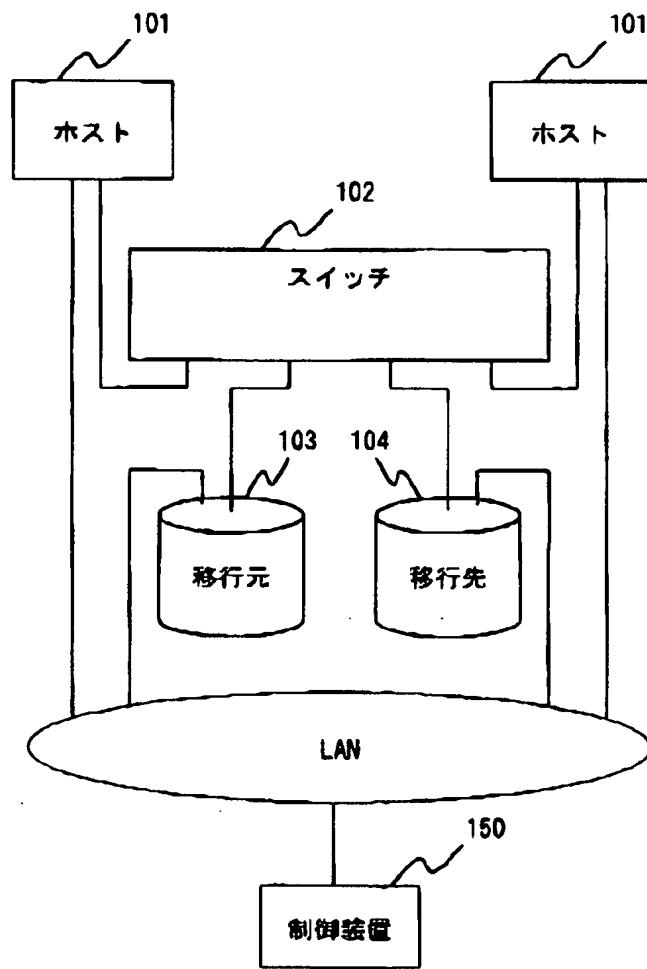


(24)

特開2001-249853

【図22】

図22



フロントページの続き

(51) Int. Cl.
G 11 B 20/10

識別記号

F I
G 11 B 20/10

テクニカル (参考)
D

F ターム(参考) SB014 EA05 EB05 GA25 GC07 GE04
HA05 HB27 HC05 HC13
SB018 GA04 HA04 MA12 QA20
SB065 BA01 CA11 CC03 EA33 PA06
ZA02 ZA03
SB082 CA10
SD044 AB01 CC04 DE49 HL01 HL06
HL11 JJ06 JJ07